

APPLICANT: Laligudi, Raghunath V.
APPLICANT: Kamigaki, Laura Y. (lco)
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
FILE REFERENCE: PL-0012-1 CON
CURRENT APPLICATION NUMBER: US/09/923,876
CURRENT FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/298,329
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/085,331
PRIOR FILING DATE: 1998-05-05
NUMBER OF SEQ ID NOS: 6332
SOFTWARE: PERL Program
SEQ ID NO 2923
LENGTH: 246
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020013958A1 700161403H1
NAME/KEY: unsure
LOCATION: 187
OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-2923

Query Match 9.2%: Score 29.6; DB 9; Length 246;

Best Local Similarity 56.0%: Pred. No. 5.7;

Matches 56; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 200 CACCAAGTGTGCTGATACGATGCGATTATTCCTTCTCTCTAGAAATGTTCTG 259
|| |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 167 CATGAAGAAGCGCGCGCTAAATTTCTTCTTGTGTTCCCTGCTGTATGCCCTCAAG 108
|| |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 260 CCGATGCTTATAGAGAGAGTGGTCAGCATGATCTCT 299
|| |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 107 AATTGCTTATTGAGATGCTTGGTCAGTACCTTCTCT 68
|| |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: August 18, 2003, 14:33:48
Job time : 164 secs

1

```

RESULT 12
US-09-766-199-42
: Sequence 42, Application US/09766399
: Patent No. US20010047092A1
: GENERAL INFORMATION:
: APPLICANT: Bruce, Wesley B.
: APPLICANT: Nlu, Xiping
: TITLE OR INVENTION: No. US20010047092A1 Plant Promoters and Methods of Use
: FILE REFERENCE: 1165
: CURRENT APPLICATION NUMBER: US/09/766,399
: CURRENT FILING DATE: 2001-01-19
: PRIOR APPLICATION NUMBER: US 60/177,437
: PRIOR FILING DATE: 2000-01-21
: NUMBER OF SEQ ID NOS: 72
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 42
: LENGTH: 30
: TYPE: DNA
: ORGANISM: Zea mays
: US-09-766-199-42

```

```

Query Match          9 3%; Score 30;   PA 9;   Length 30;
Best Local Similarity 100.0%; Pred. No. 1 4;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      155  AATCTACCTCCACACCAACCCAGCTTGTGA 184
          |||||
DB       1  AATCTACCTCCACCAACCAACGAGCTTTGTA 30

RESULT 13
US-09-766-199-43
; Sequence 43, Application PIS/09766399
; Patent No. US20010047092A1
; GENERAL INFORMATION:

```

```
QY      106 ATCACCACCACTATCACCTAGAAAAGCGA   135
          |||||
Db       1 ATCACACCAACTTATCACCTAGAAAAGCGA   30
```

```

RESULT 14
US-10-050-704-57/c
: Sequence 57, Application US/10u50704
: Publication No. US20030050442A1
GENERAL INFORMATION:
APPLICANT: Ruden et al.
TITLE OF INVENTION: 62 Human Secreted Proteins
FILE REFERENCE: P2039P1
CURRENT APPLICATION NUMBER: US/10/050,704
CURRENT FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: 09/684,524
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: PCT/US00/08979
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128,693
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 60/130,991
PRIOR FILING DATE: 1999-04-26
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 57
LENGTH: 1769
TYPE: DNA
ORGANISM: Homo sapiens
US-10-050-704-57

```

	Query Match	9.38%	Score 29.8	DB 14	Length 1769
	Best Local Similarity	51.1%	Pred. No. 13		
	Matches	70	Conservative	0	Mismatches 67
					Indels 0
					Caps 0
QY	152	TGCAATCTACCTGCACCAACCCAGCTTGTATCTCTACTGTGATCACCACAAATCT	211		
Db	1395	TGCAATCTCCCTGGTGCCTCTTATCTCTGGATTTCTATCTCTCTCTCTCTCTCTCTCT	1336		
QY	212	CGTGATACCATGTGCGCATTTATTCCTCTTCTCTCTACAAATCTTCTGCCCATCTTTAT	271		
Db	1335	GGAGCTACCATGCTCCCTGCTGCGCTTATTTAAAAAGAAACAGTCCCGCGCTTTCTCT	1276		
QY	272	AAGAGAAGTTGGTCTAG	288		
Db	1275	TGGTTGTTGCTGCTCCAG	1259		

```

RESULT 15
US-09-923-876-2923/c
; Sequence 2923, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:

```


Query Match	100.0%	Score 322	DB 13	Length 322
Best Local Similarity	100.0%	Proc. 322	1.2e-101	
Matches 322	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	GAATTCCTGCTCTGGGGGTCAACTGACCGGTAAACAGTGAAGATGATATCTCTTCTCT	60	
Db	1	GAATTTCTGCTCTCGCCGGCTCAACTGACCGGTAAACAGTGAAGATGATCTCTTCTCT	60	
QY	61	CTCTGCAATCCGTCGCGCTGGAGCAAAATGGGGCAGTCCGCTACCTATCAACACCAACTTAT	120	
Db	61	CTCTGCAATCCGTCGCGCTGGAGCAAAATGGGGCAGTCCGCTACCTATCAACACCAACTTAT	120	
QY	121	CACCTAGAAAAGGAGACGGCTCTCTGCAATGCAATTCACCTCCCAACCAACCCACACTT	180	
Db	121	CACCTAGAAAAGGAGACGGCTCTCTGCAATTCACCTCCCAACCAACCCACCACTT	180	
QY	181	TGATCTGCTTACTGATGATCAACCAAGATTGCTGATACATATGATATATTTCTCTT	240	

RESULT 4
 US-08-987-367-3/c
 Sequence 3, Application US/08987367
 Patent No. 6399859
 GENERAL INFORMATION:
 APPLICANT: Nichols, Scott
 APPLICANT: Dhugga, Kanwarpal S.
 APPLICANT: Singletary, George William
 APPLICANT: Saunders, Court
 APPLICANT: Wilcher, Derrick
 APPLICANT: Bruce, Wesley B.
 APPLICANT: Sims, Lynne
 APPLICANT: Lu, Guihua
 APPLICANT: Zhong, Gan Yuan
 TITLE OF INVENTION: PLANT URIDINE DIPHOSPHATE-GLUCOSE
 NUMBER OF INVENTION: DEHYDROGENASE GENES, PROTEINS, AND USES THEREOF
 NUMBER OF SPOUNCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED AND BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

1 APPLICATION NUMBER: US-08/987,367
 2 FILING DATE: 10-DEC-1997
 3 CLASSIFICATION:
 4 ATTORNEY/AGENT INFORMATION:
 5 NAME: Jones, Phil B.C.
 6 REGISTRATION NUMBER: 38,195
 7 REFERENCE/DOCKET NUMBER: 750027, 404
 8 TELECOMMUNICATION INFORMATION:
 9 TELEPHONE: (206) 622-4900
 10 TELEFAX: (206) 682-6031
 11 INFORMATION FOR SEQ ID NO: 3:
 12 SEQUENCE CHARACTERISTICS:
 13 LENGTH: 1947 base pairs
 14 TYPE: nucleic acid
 15 STRANDEDNESS: single
 16 TOPOLOGY: linear
 17 US-08-987-367-3

Query Match	9.9%	Score 31.8;	DB 4;	Length 1947;
Best Local Similarity	67.2%;	Pred. No. 0.41;		
Matches 45; Conservative	0;	Mismatches 22;	Indels 0;	Gaps 0;

0y 14 GGGCGTCACTGAACCGTAACAGTGATGACTTTTCCTCTGCAATCCGT 73
||||| | | | | | | | | |
Db 113 GGGGTCAACCGCACCGAGGGAGAAGAGGAGCGGCATCTCTGCGCTTCCT 54

QY	74	GCCGTGC	80
Db	53	GGCCCGC	47

RESULT 5
US-09-369-618-3
; Sequence 3, Application US/09369618
Date: 11/10/2011

```

1 GENERAL INFORMATION:
2 APPLICANT: Lok, Si
3 TITLE OF INVENTION: HUMAN PROHORMONE CONVERTASE 4
4 FILE REFERENCE: 97-0502
5 CURRENT APPLICATION NUMBER: US/09/369,618
6 CURRENT FILING DATE: 1999-08-06
7 EARLIER APPLICATION NUMBER: US 09/071,101
8 EARLIER FILING DATE: 1998-05-01
9 EARLIER APPLICATION NUMBER: US 60/044,015
10 EARLIER FILING DATE: 1998-05-06
11 NUMBER OF SEQ ID NOS: 16
12 SOFTWARE: FASTSEQ for Windows Version 3.0
13 SEQ ID NO 3
14 LENGTH: 2265
15 TYPE: DNA
16 ORGANISM: Artificial Sequence
17 FEATURE:
18 OTHER INFORMATION: Degenerate polynucleotide sequence for human
19 FEATURE:
20 OTHER INFORMATION: prohormone convertase 4
21 FEATURE:
22 NAME/KEY: misc_feature
23 LOCATION: (1)..(2265)
24 OTHER INFORMATION: n = A,T,C or G
25 US-09-369-618-3

```

Query Match 9 68; Score 31; DB 3; Length 2265;
Best Local Similarity 26 48; Pred. NO. 0.84;
Matches 73; Conservative 42; Mismatches 159, Indels 2, Gaps 1

Dy 11 CTCGGCGGCTAACTGAACGTAAATAGTGAAGAAGGATACTTTTCCTCCTCTGCATTC 70
| | | | : : : : | ||:
Db 1913 CNGNCAYACAGCCGCCNCCNGCANTMGNCTNTGTWSSNNSTGYCAYGCMMSNTGYTAYA 1972

Qy 71 CGTCCCGGTGGGAAGCAATATG6GCCGACGTGCCTACTTATCACACCAACTTATCACCCTAGAAA 130
+ + + + + + + + + + + + + + + + + +
Db 1973 CATTGYMGNGNGGAMNSNCCNMGNCAVTTATACNWSNTGTTCCTCCMMSNMSNAACNTYTNATYC 2033

[illegible]

RESULT 6
US-09-369-617-3

```

; Sequence 3, Application US/09369617
; Patent No. 6127162
; GENERAL INFORMATION:
;

```

```

: APPLICANT: Lok, SI
: APPLICANT: Jaspers, Stephen R.
: TITLE OF INVENTION: HUMAN PROHORMONE CONVERTASE 4
: NUMBER OF PAGES: 1

```

;; CURRENT APPLICATION NUMBER: US/09/369,617
;; CURRENT FILING DATE: 1999-08-06
;; EARLIER APPLICATION NUMBER: US 09/071,101

EARLIER APPLICATION NUMBER: US 60/044,015
EARLIER FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 16

```

; LENGTH: 2265
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

: OTHER INFORMATION: Degenerate polynucleotide sequence for human
:
: OTHER INFORMATION: prothrombin convertase 4
:
: FEATURE:
:
: NAME/KEY: misc_feature
: LOCATION: (1) (2265)
: OTHER INFORMATION: n = A,T,C or G
: DS-09-369-617-3

```

| | | | | |
|-----------------------|------------------|-----------------|-----------|--------------|
| Query Match | 9.68; | Score 31; | DB 3; | Length 2265; |
| Best Local Similarity | 26.48; | Pred. NO. 0.84; | | |
| Matches 73; | Conservative 42; | Mismatches 159; | Indels 2; | Gaps 1; |

Gy 11 CTGAGACATCAACTGAACCTGTAAACAGTGGAAAAGTGATACCTCTTCCTCCTGCATC 70
| | | | |
Db 1913 CAGCGNCAVACNGCNGCNCNCNTTNMGNGNTGYWSWMSNGCYCAGCNMSNTGTAYA 19

Dy 71 CGTCCCGCTGGAGCAAAATGCGCGCAGTCGCCTACTTATACACCACAATTATCACCTTAGAAA 130
| | : | | : | : | |
Db 1973 CNTCTGMAKNGKNGMNSNCTNMGNALYTGTA CNMSITTGYCCNCCTMNSMWSNAQNTNGAYC 2032

Oy 131 AGGCAC--GCGTCTGGATGCGATTGCAAACTTACCTCCACCAACCCAGCTTGTATCTG 188
 | : | : : | | | | : : : | : | : |
 Ph 2033 APCACACACGCGNMSNTGYATGGGCGCCNACNACGCGCGATWSNMCGMCCNMGCTTTNCGCAG 2092

OY 189 CTTACTGTATCATCAAAATTCTTCTGATACAGATCTGCATTAAGCTCTTTCTTCCTCA 248
| | | | | | | | | | : : : :
Dp 2093 GCGTGTTGCCNCACATCAAGTGTGCCGCGMSNGCAGATGGTNTYTWSNNTNYTNCGTGA 2153

Dy 249 GAATGTTCTCGCATGCTTAAAGAGGAACG 284
| | : | : | :
Db 2153 CNYTCGGCNGCCNGTYNTNCTGCNAATGWSNATCG 2188

```

RESULT 7
US-09-313-294A-5664/C
; Sequence 5664, Application: US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:

```

```

RESULT 7
US-09-313-294A-5664/C
; Sequence 5664, Application: US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:

```

LENGTH: 1830121 base pairs

[illegible]

19 TCACTGACCTTAAACAGTGGAAAGTATCTTCTCTCTGCAATCCGCTGCT 78
1401922 TTACTTAATTCCTTAACAGTCCGGAAGGAAATATTACCTGACCAAGACCTACTT 1401864
79 GGAAGCAATATGAGTATGCTGCTTATATATACCAACTATACCTAGAAAGGAGG 137
1401863 GCAAGCAATATGAGTATGCTGCTTATATATACCAACTATACCTAGAAAGGAGG 1401805

RESULT 10
US-09-328-352-3148/C
Sequence 3148, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328.352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 3148
LENGTH: 699
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-3148

Query Match 9.1%; Score 29.2; DB 4; Length 699;
Best Local Similarity 56.1%; Pred. No. 2.1;
Matches 55; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

224 TCGCATTTATGCTCTCTCTCTGCAATGCTTCTGCGGATCTTTAAGAGAGCTTG 283
576 TACATTAATTAAGACCTTTTTCATGCTGCTGTTATGAACGCTTTCAATAAGATGTCG 517
284 GTCAGCATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
516 AATGCTGATGATGCTTATTAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 479

RESULT 11
US-09-197-063-1/C
Sequence 1, Application US/09197063
Patent No. 6261817
GENERAL INFORMATION:
APPLICANT: Zalacatin, Magdalena
APPLICANT: Brown, James R.
APPLICANT: Biswas, Sanjoy
APPLICANT: Warren, Richard L.
APPLICANT: Shilling, Lisa K.
TITLE OF INVENTION: No. 6261817el Cuda
FILE REFERENCE: GM10121
CURRENT APPLICATION NUMBER: US/09/197.063
CURRENT FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: 60/066,350
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1551
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: CDS
LOCATION (1) (1548)
US-09-197-063-1

Query Match 9.1%; Score 29.2; DB 3; Length 1551;
Best Local Similarity 52.5%; Pred. No. 3;
Matches 64; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
61 CTCTGCAATCCGTGCTGGAAGCAATGGCGAGTGGCTACTTATACACCAACTTAT 120

1330 CAGTCTTAAGCATGATGATGCAATATGCGGCTGATCCAGCTTTAGCATTTCT 1271
121 CACCTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
1270 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1211
181 TG 182
1210 TG 1209

RESULT 12
US-09-221-294-3
Sequence 3, Application US/09221294
Patent No. 626138
GENERAL INFORMATION:
APPLICANT: Riccardo Dalla-Favera and
Alessandro Massimo Gianni
TITLE OF INVENTION: A Retroviral Vector Capable of Transducing the
TITLE OF INVENTION: Aldehyde Dehydrogenase-1 Gene and Uses of Said
TITLE OF INVENTION: Vector
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM 330 466 DX2
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221.294
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 42990-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-391-0525
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2904 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: N
ANTI-SENSE: N
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 363..2274
OTHER INFORMATION:
US-09-221-294-3

Query Match 9.1%; Score 29.2; DB 3; Length 2904;
Best Local Similarity 53.5%; Pred. No. 4;
Matches 61; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

26 AACGCTAAAGAGTGGAAAGTATGATGCTTCTCTCTGCAATCCGCTGCGGAGAGCA 85
844 AACCCAGCCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 903
86 AATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 139
904 AATGCTAAAGAGTGGAAAGTATGATGCTTCTCTCTGCAATCCGCTGCGGAGAGGAG 957

Search completed: August 18, 2003, 14:30:51
Job time : 63 secs

Search completed: August 18, 2003, 14:30:51
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 18, 2003, 13:14:55 ; Search time 1913 seconds
(without alignments)
4090.978 Million cell updates/sec

Title: US-10-058-566-5

Perfect score: 322

Sequence: 1 gaattcgtctgcgcggtc.....agtgtctagctgagaacatg 322

Scoring table: IDENTITY-NWG

Gapop 10 0 ; Gapext 1 0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estimu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hic:*
9: qb_est1:*
10: qb_est2:*
11: qb_hic:*
12: qb_est3:*
13: qb_est4:*
14: qb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_tod:*
26: em_gss_ping:*
27: em_gss_vrt1:*
28: qb_gss1:*
29: qb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|-------|-------------|
| 1 | 223.8 | 69.5 | 740 | 29 CC349513 |
| 2 | 190.8 | 59.3 | 508 | 29 B2401831 |
| 3 | 190.8 | 59.3 | 524 | 29 B2401828 |
| 4 | 45.8 | 14.2 | 685 | 12 B211851 |

| | | | | | |
|------|------|------|------|----|----------|
| C 5 | 38.2 | 11.9 | 533 | 9 | AW784857 |
| C 6 | 37 | 11.5 | 849 | 13 | B0745750 |
| C 7 | 36.8 | 11.4 | 664 | 10 | BB449721 |
| C 8 | 36.8 | 11.4 | 675 | 9 | AV349139 |
| C 9 | 36.8 | 11.4 | 2560 | 11 | AK032641 |
| C 10 | 36.8 | 11.4 | 2574 | 11 | AK051108 |
| C 11 | 36 | 11.2 | 1053 | 13 | B0745776 |
| C 12 | 36 | 11.2 | 1069 | 13 | B0745749 |
| C 13 | 35.6 | 11.1 | 549 | 28 | AA799378 |
| C 14 | 35.4 | 11.0 | 607 | 12 | B0543444 |
| C 15 | 35.4 | 11.0 | 612 | 12 | B0101888 |
| C 16 | 35.4 | 11.0 | 627 | 12 | B0101944 |
| C 17 | 35.2 | 10.9 | 588 | 10 | BC524736 |
| C 18 | 34.8 | 10.8 | 935 | 29 | CNS03A3N |
| C 19 | 34.6 | 10.7 | 522 | 12 | B3501394 |
| C 20 | 34.4 | 10.7 | 284 | 10 | BB245583 |
| C 21 | 34.4 | 10.7 | 665 | 12 | B0529082 |
| C 22 | 34.4 | 10.7 | 803 | 13 | B0745777 |
| C 23 | 34 | 10.6 | 385 | 12 | BM798289 |
| C 24 | 33.8 | 10.5 | 859 | 28 | BM655718 |
| C 25 | 33.6 | 10.4 | 265 | 9 | AA348378 |
| C 26 | 33.6 | 10.4 | 671 | 28 | BB453718 |
| C 27 | 33.6 | 10.4 | 800 | 28 | BB488995 |
| C 28 | 33.6 | 10.4 | 960 | 29 | CNS080PA |
| C 29 | 33.4 | 10.4 | 608 | 12 | B026852 |
| C 30 | 33.4 | 10.4 | 1201 | 29 | CNS016CY |
| C 31 | 33.2 | 10.3 | 348 | 14 | T01667 |
| C 32 | 33.2 | 10.3 | 482 | 9 | AM653990 |
| C 33 | 33.2 | 10.3 | 685 | 28 | BB700981 |
| C 34 | 33.2 | 10.3 | 751 | 28 | BB461211 |
| C 35 | 33 | 10.2 | 408 | 14 | T18714 |
| C 36 | 33 | 10.2 | 537 | 28 | AO765571 |
| C 37 | 33 | 10.2 | 540 | 14 | CB144345 |
| C 38 | 33 | 10.2 | 1192 | 12 | BB787172 |
| C 39 | 33 | 10.2 | 1201 | 9 | AL552063 |
| C 40 | 32.8 | 10.2 | 453 | 9 | AA941154 |
| C 41 | 32.8 | 10.2 | 632 | 29 | B2905637 |
| C 42 | 32.8 | 10.2 | 724 | 28 | BB695201 |
| C 43 | 32.8 | 10.2 | 742 | 29 | AG085810 |
| C 44 | 32.8 | 10.2 | 768 | 28 | BB547822 |
| C 45 | 32.8 | 10.2 | 899 | 29 | CNS04200 |

ALIGNMENTS

RESULT 1
LOCUS CC349513 740 bp DNA linear GSS 16-MAY-2003
DEFINITION CC00AS63TH_ZM_0.7.1.5_KB Zea mays genomic clone ZMWB0368105,
genomic survey sequence.
ACCESSION CC349513
VERSION CC349513.1 GI:30818920
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta; Magnoliophyta; Liliopsida, Poales, Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 740)
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick
, A., Fraser, C.M., Budiman, M.A., Rodell, J.A., Rohlfing, T., Citek
, P.W., Nunnberg, A., Pobbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished
Contact: Cathy Whitelaw
TICK 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@ligr.org
Seq primer: TR
Class: sheared ends.

```

FEATURES
  source
    location/Qualifiers
    1..740
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBma0015620"
    /note="Vector: pBCSK-; Site: 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

BASE COUNT
  187 a 184 c 164 g 205 t

Query Match
  Best Local Similarity 85.6%, Pred. No. 8.5e 59;
  Matches 273; Conservative 0; Mismatches 42; Indels 4; Gaps 2;

OY 6 CTGCTCCGGCGGTCAACGTGAACCGTAACAGTGCATCTCTTCTCTCTG 65
    |||||
DB 1 CTGCTCCGGCGGTCAACGTGAACCGTAACAGTGCATCTCTTCTCTCTG 60
    |||||

OY 66 CAATCCGTCGCGTGAAGCAATGCGCGCTACTATACACCACTTATCACT 125
    |||||
DB 61 CCGT--AAAGCGGTCAACTACACCGGTAGCATTTGTTGACACCACTTATCACT 118
    |||||

OY 126 AGAAAGCGACGCGTCTGATGATTCGAATCTACCTCAACCAACCCACTTTGAT 185
    |||||
DB 129 ACAAAGCGACGCGTCTGATGATTCGAATCTACCTCAACCAACCCACTTTGAT 178
    |||||

OY 186 CTGCTTACTGTATCAACCAAGTTGCTGATACGATGCGCATTTGCTTCTCTCT 245
    |||||
DB 179 CTGCTTACTGTATCAACCAAGTTGCTGATACGATGCGCATTTGCTTCTCTCT 238
    |||||

OY 246 CTAGAAATGTTCTGCGCATGCTTATAGAGAAAGTGTGTCAGC--ATCGATCTGCGCA 303
    |||||
DB 239 CTAGAAATGTTCTGCGCATGCTTATAGAGAAAGTGTGTCAGCATATGATCTGCGCA 298
    |||||

OY 304 CTGCTTACTGTATCAACCAAGTTGCTGATACGATGCGCATTTGCTTCTCTCT 322
    |||||
DB 299 GTGCTAGTGTAGAACATG 317
    |||||

RESULT 2
LOCUS BZ401831 508 bp DNA linear GSS 04-DEC-2002
DEFINITION OGAAG517M ZM_0.7-1.5_KB Zea mays genomic clone ZMMBma0015620,
ACCESSION BZ401831
VERSION BZ401831.1 GI:26026901
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 508)
AUTHORS White, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick
  , A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek
  , R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished
  Other GSSs: OGAAG517C
  Contact: Cathy White, Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitec@tifgr.org
  Seg primer: TR
  Class: sheared ends.
  Location/Qualifiers
  1..508
  /organism="Zea mays"

FEATURES
  source
    location/Qualifiers
    1..524
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBma0015620"
    /note="Vector: pBCSK-; Site: 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

BASE COUNT
  135 a 129 c 107 g 137 t

Query Match
  Best Local Similarity 83.8%, Pred. No. 1.5e 48;
  Matches 228; Conservative 0; Mismatches 42; Indels 2; Gaps 1;

OY 6 CTGCTCCGGCGGTCAACGTGAACCGTAACAGTGCATCTCTTCTCTCTG 65
    |||||
DB 239 CTGCTCCGGCGGTCAACGTGAACCGTAACAGTGCATCTCTTCTCTCTG 298
    |||||

OY 66 CAATCCGTCGCGTGAAGCAATGCGCGCTACTATACACCACTTATCACT 125
    |||||
DB 299 CCGT--AAAGCGGTCAACTACACCGGTAGCATTTGTTGACACCACTTATCACT 356
    |||||

OY 126 AGAAAGCGACGCGTCTGATGATTCGAATCTACCTCAACCAACCCACTTTGAT 185
    |||||
DB 357 ACAAAGCGACGCGTCTGATGATTCGAATCTACCTCAACCAACCCACTTTGAT 416
    |||||

OY 186 CTGCTTACTGTATCAACCAAGTTGCTGATACGATGCGCATTTGCTTCTCTCT 245
    |||||
DB 417 CTGCTTACTGTATCAACCAAGTTGCTGATACGATGCGCATTTGCTTCTCTCT 476
    |||||

OY 246 CTAGAAATGTTCTGCGCATGCTTATAGAGAAAGTGTGTCAGC--ATCGATCTGCGCA 277
    |||||
DB 477 CTAGAAATGTTCTGCGCATGCTTATAGAGAAAGTGTGTCAGCATATGATCTGCGCA 508
    |||||

RESULT 3
LOCUS BZ401828/c 524 bp DNA linear GSS 04-DEC-2002
DEFINITION OGAAG517C ZM_0.7-1.5_KB Zea mays genomic clone ZMMBma0015620,
ACCESSION BZ401828
VERSION BZ401828.1 GI:26026898
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 524)
AUTHORS White, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick
  , A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek
  , R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished
  Other GSSs: OGAAG517M
  Contact: Cathy White, Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitec@tifgr.org
  Seg primer: TR
  Class: sheared ends.
  Location/Qualifiers
  1..524
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /strain="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBma0015620"
  /note="Vector: pBCSK-; Site: 1: HincII; 0.7-1.5 kb
  methylation filtered genomic DNA library"

```


| | | |
|----------------------------------|---|--|
| REFERENCE | 1 | Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning with a cDNA library. <i>Methods Enzymol.</i> 303, 19-44 (1999) |
| REFERENCE | 2 | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. <i>Genome Res.</i> 10 (10), 1617-1630 (2000) |
| JOURNAL MEDLINE PUBMED REFERENCE | 3 | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsuno, H., Sadaquchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watanabe, T., Onoda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. <i>Genome Res.</i> 10 (11), 1757-1771 (2000) |
| JOURNAL MEDLINE PUBMED REFERENCE | 4 | Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Atsawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Reichmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stamboli, F., Suzuki, K., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Boljuga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, C., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schobach, C., Seyer, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Wetz, C., Whitaker, C., Wilmberg, L., Wyszynski, B., Yamaoka, Y., Yoshida, K., Hasegawa, Y., Kawai, H., Kotsuki, S., and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection. <i>Nature</i> 409 (6821), 685-690 (2001) |
| JOURNAL MEDLINE PUBMED REFERENCE | 5 | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. <i>Nature</i> 420, 563-573 (2002) |
| JOURNAL MEDLINE PUBMED REFERENCE | 6 | (Bases 1 to 2560) |
| JOURNAL MEDLINE PUBMED REFERENCE | 7 | Adachi, J., Aizawa, K., Akimura, T., Atsawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hangaki, T., Hara, A., Hashizume, W., Hayashi, K., Hayatsu, N., Hiramoto, K., Hltaka, T., Hirozane, T., Horii, F., Imoto, K., Ishii, Y., Itoh, M., Kageawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohgato, N., Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, I., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takekida, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission |
| JOURNAL MEDLINE PUBMED REFERENCE | 8 | Submitted (16-Jun-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC). |

| | COMMENT | FEATURES |
|------------|--|---|
| | <p>Riken Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama Kanagawa 230-0045, Japan (E-mail: genome-res@gsic.riken.go.jp/, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)</p> <p>cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.</p> <p>Please visit our web site for further details.</p> <p>URL: http://genome.gsc.riken.go.jp/
URL: http://antom.gsc.riken.go.jp/.
Location/Qualifiers
1..2560</p> | <p>SOURCE</p> <p>/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6530401M22"
/db_xref="taxon:10090"
/clone="6530401M22"
/issue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
413..1193
/note="unnamed protein product; putative similar to HYPOTHETICAL 34.3 KDA PROTEIN [Homo sapiens] (SPTRIO96BM1, evidence: FASTA, 82.6%ID, 100%length, match=978) "
/codon_start=1
/protein_id="BAC27966.1"
/db_xref="GI:2632845"
/translation="MPWDTGRSGANGSPGPAARLRVQKCRKSFAFYLAVRDQI
PVLLEILRASAFPCDEGRRAAYSPSEALYLAVHDQAVAHYLLATFPICALAPLPI
SAGFCETARPCHVALAVARYRVGLIRLIITVDPEPERVRLIDRSCSVGGCTG
SHLVACELAPCEFLLEIGHGASPLGDLGGSPFLLEILLRNQNPASSAPTAAEAS
TNNANNTSTSEEVCORRILLIDLVLTYPGGVNVPAREILIGDLPBORLGSDRK
QMLAGLAPSLFVRMOVLYTTISRGKFFPALDELPLPSFLQPLDLTGNG"</p> |
| CDS | | |
| | <p>Query Match
Best Local Similarity 54.4%; Pred. No. 4.6;
Matches 74; Conservative 0; Mismatches 62; Indels 0; Gaps 0;</p> | |
| BASE COUNT | 453 a 744 c 804 g 359 t | |
| ORIGIN | | |
| Db | 2334 CTCTGTCTCTGCACATCATCGCCGTCGAAGCCTATTGCCAAGAAGTTGAATTGCACA 2275 | |
| QY | 51 CTCTTCTCTCTGCATTCGTCGCGCTGGGAAGCAAAATGGCCGATCGCTACTATTAC 110 | |
| Db | 111 ACCAATTATCACTAGAAAAGCAGCGGTCTTGATTCGATTGCAAAATTACTCTCAAC 170 | |
| QY | 2274 GCCCACAAATATCCAGGATAGGAGCGGCCCTCCCGCACTTCMAAGTTCCCTGGTATG 2215 | |
| Db | 171 AACCCAGCTTTGTATC 186 | |
| QY | 2214 AACCCAGCTTTGCAGC 2199 | |
| RESULT 10 | | |
| AKO51108/c | AKO51108 | 2574 bp mRNA linear HTC 05-DEC-2001 |
| Locus | Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched library, clone:ID030073E24 product:similar to HYPOTHETICAL 34.3 KDA PROTEIN [Homo sapiens], full insert sequence. | |
| Accession | AKO51108 | |
| Version | AKO51108.1 | GI:26341731 |
| Keywords | HTC; CAP trapper. | |
| Source | Mus musculus (house mouse) | |
| Organism | Mus musculus | |
| Reference | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | |
| Authors | Carninci, P. and Hayashizaki, Y. | |
| Title | High-efficiency full-length cDNA cloning | |

1 1053

BASE COUNT
ORIGIN

Host Local
Matches 5

10b

Db

DEFINITION
ACCESSION

KEYWORDS
SOURCE

REFERENCES

**JOURNAL
COMMENT**

Tel: 615 936 2660
Fax: 615 936 2661

High quality sequence start: 59
High quality sequence stop: 826

| FEATURES | SOURCE |
|----------|--------|
|----------|--------|

BASE COUNT
ORIGIN

Best Local
Matches 5

Db

Db

DEFINITION

KEYWORDS
SOURCE

REFERENCE

JOURNAL
COMMENT

Class: plasmid ends
High quality sequence stop: 549.

FEATURES

SOURCE

```
/clone="UUGC2M0047K08"
/sex="Male"
```


/lab.host="E. coli strain XL10-gold, T1-resistant, F-"
/clone.lib="Mouse 10kb plasmid UUCGM library"
/note="Vector: pMD42ny; purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114) [pLAR129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-gold (Stratagene) cells and selected for ampicillin resistance."

| | | | | | |
|-----------------------|---|---------------|------------|------------|----------|
| BASE COUNT | 132 a | 164 c | 183 g | 127 t | 1 others |
| ORIGIN | | | | | |
| Query Match | 11.0% | Score 35.4 | DB 12 | Length 607 | |
| Best Local Similarity | 59.4% | Pred No 6.1 | | | |
| Matches | 60; Conservative | 0; Mismatches | 41; Indels | 0; Gaps | 0; |
| UY | 188 GTTATGTCAGTCAAAAATTGTCTCTCATACGACATCGAATATATGCTTTCTTGCTT 247 | | | | |
| | | | | | |
| Db | 1 GTTAATGGATGACAAGAAGGCTTGACACACACAAAGCCGCTCCTACCTCTTCCTCTCT 60 | | | | |
| | | | | | |
| UY | 248 AATAATGCTCTCCGACATGCTTATGAAGAGAGTTGAGTAG 288 | | | | |
| | | | | | |
| Db | 61 GATATGTCCTGCATATGACATTAATTTAGAGAGTGGAGAG 101 | | | | |
| | | | | | |
| RESULT 15 | | | | | |
| BJ012888/c | 612 bp MPNA linear EST 05-DEC-2001 | | | | |
| DEFINITION | BJ012888 MF01SSA CDNA Oryzias latipes cdna clone MF01SSA181A11 5' | | | | |
| LUCUS | mRNA sequence. | | | | |
| ACCSSION | BJ012888 GI:17363321 | | | | |
| VERSION | EST. | | | | |
| KEYWORDS | Oryzias latipes (Japanese medaka) | | | | |
| SOURCE | Oryzias latipes | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Nuteleostomi; Fkariyotae; Teleostei; Euteleostei; Euteleostei; Euteleostei; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Acanthopterygii; Perciformes; Atherinomorpha; Atherinomorpha; Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias. | | | | |
| REFERENCE | 1 (bases 1 to 612) | | | | |
| AUTHORS | Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jinbo,T. and Takeda,H. | | | | |
| TITLE | Medaka EST Project in Takeda's lab | | | | |
| JOURNAL | Unpublished | | | | |
| COMMENT | Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tsuhii@genes.nig.ac.jp. | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..612 | | | | |
| | /organism="Oryzias latipes" | | | | |
| | /mol_type="mRNA" | | | | |
| | /strain="Md-ir" | | | | |
| | /db_xref="taxon:8090" | | | | |
| | /clone="MF01SSA181A11" | | | | |
| | /sex="mixture of female and male" | | | | |
| | /tissue_type="whole embryo" | | | | |
| | /dev_stage="segmentation stage 20 - 25" | | | | |
| | /clone_idb="MF01SSA CDNA" | | | | |
| BASE COUNT | 126 a | 183 c | 165 g | 138 t | |
| ORIGIN | | | | | |
| Query Match | 11.0% | Score 35.4 | DB 12 | Length 612 | |
| Best Local Similarity | 59.4% | Pred No 6.1 | | | |
| Matches | 60; Conservative | 0; Mismatches | 41; Indels | 0; Gaps | 0; |
| CY | 188 GTTATGTCAGTCAAAAATTGTCTCTCATACGACATCGAATATATGCTTTCTTGCTT 247 | | | | |
| | | | | | |
| Db | 602 GTTAATGGATGACAAGAAGGCTTGACACACACAAAGCCGCTCCTACCTCTTCCTCTCT 543 | | | | |
| | | | | | |
| OY | 248 AGAATGCTCTCCGCCATGCTTATTAACGAAAGCTGCTGAC 288 | | | | |
| | | | | | |
| Db | 542 GAATGCTCTCTCATATGACATTAATTTAGAGAGTGGAGAG 502 | | | | |
| | | | | | |

Search completed: August 18, 2003, 14:29:32
Job time : 1918 secs

.

1

[illegible]

XX (PION-) PIONEER HI-BRED INT INC.
PA

XX Albertsen M, Fox T, Hufman G, Trimmel M;
 XX
 XX WPI: 2002-657541/70.
 DR
 XX
 XX New isolated DNA molecule, useful for mediating male fertility in maize
 PT plants to produce a hybrid that may increase the yield of desirable
 PT characteristics in plants, e.g. resistance to disease and insects or
 PT heat and drought tolerance
 XX
 PS Claim 38: Fig 10; 54pp; English.
 XX
 CC The invention relates to a DNA molecule mediating male fertility in
 CC plants comprising the BS92-7 gene, represented by American Type Culture
 CC Collection (ATCC) deposit number 98932. The DNA molecules are useful in
 CC mediating male fertility in plants, particularly maize, to produce a
 CC hybrid plant that may increase the yield and combination of desirable
 CC characteristics in plants, such as resistance to disease and insects or
 CC heat and drought tolerance. The present sequence represents the maize
 CC BS92-7 promoter essential region sequence.
 CC
 SO Sequence 187 BP; 42 A; 46 C; 40 G; 59 T; 0 other;

Query Match 58.1%; Score 187; DB 24; Length 187;
 Best Local Similarity 100.0%; Pred. No. 1.9e-54;
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 CCGGCTCGATGATGATGCAATGCTTACTCCACCAACCCAGCTTGTATCTGCTTCTG 195
 DB 1 CCGGCTCGATGATGATGCAATGCTTACTCCACCAACCCAGCTTGTATCTGCTTCTG 60
 QY 196 TGATCAGCAAGTGTGCTGATAGCATGTGCGATTTGCTTCTCTCTAGATGTT 255
 DB 61 TCATCACCAGTGTGCTGATAGCATGTGCGATTTGCTTCTCTCTAGATGTT 120
 QY 256 CCGGCGATGCTTTTAAAGAGAGTGTGTAATGATTTTGAATGATTTCTAGCTGA 315
 DB 121 CCGGCGATGCTTTTAAAGAGAGTGTGTAATGATTTTGAATGATTTCTAGCTGA 180
 QY 316 GACATG 322
 DB 181 GACATG 187

RESULT 4
 ABL10980/C
 ID ABL10980 standard; cDNA; 3413 BP.
 XX
 AC ABL10980;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 27422.
 XX
 KM Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PH CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 PT
 DR WPI: 2001-656860/75.

DR P-PSDB; ABB66877.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 XX
 PS Claim 1: SEQ ID NO 47422; 41pp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB157737-AB172072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://pub.int/published_pcl_sequences.
 CC
 SO Sequence 3413 BP; 997 A; 759 C; 673 G; 984 T; 0 other;

Query Match 11.1%; Score 35.6; DB 23; Length 3413;
 Best Local Similarity 51.2%; Pred. No. 0.19;
 Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 154 CAATCTACCTCCACCAACCCAGCTTGTATCTGCTTACTGATACCAACTGTGC 213
 DB 402 CATTTCTACTTAAAGCGACGACAGTAATATCCCTTCATGCTCTCATTCACAGATGTA 343
 QY 214 TGATAGCATGTGCGATTTATGCTTCTTCTTCTTGAATGTTTCTGCGATGCTTATA 273
 DB 342 TCACGCAAAATGCCAATATGAGCTTCTATCAAAATTTTGGCGGACCGCTATA 283
 QY 274 GAGCAAGTTCGTCGACATGATCTCTGCACTGTCTAGCTGA 315
 DB 282 GCGATGTTGGATGTTCTATCTGCGAGTTTACTGCTTTGAGGA 241

RESULT 5
 ABL12129/C
 ID ABL12129 standard; cDNA; 4329 BP.
 XX
 AC ABL12129;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 30869.
 XX
 KM Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PH CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 PT
 DR WPI: 2001-656860/75.
 DR P-PSDB; ABB68026.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Claim 1; SEQ ID NO 30866; 21bp + Sequence Listing; English.
 PS (ABB57737-ABB72072).
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX Sequence 4329 BP; 1231 A; 1000 C; 1027 G; 1071 T; 0 other;
 SO
 Query Match 11.0%; Score 35.4; DB 23; Length 4329;
 Best Local Similarity 57.8%; Pred. No. 0.25;
 Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 Oy 43 AGTGAATACCTCTTCTCTCTGCAATCCGCGCGGAGCAAGCAATGGCGCAGTGGCTTA 102
 Db 2644 ATTGGAACCTTTCTTCTCTGCAATCCGCGCGGAGCAATGGCGCAGTGGCTT 2585
 Oy 103 CTTATATACCACTTATATATAGAAAGGAGCGCTCTGGATGCAT 151
 Db 2584 CGTCGCGCAAACTTATATCTAAATACAGCAATGCGCTCGATTAAT 2536
 RESULT 6
 AB12128
 ID AB1.12128 standard; cDNA; 9404 BP.
 AC AB1.12128;
 XX 26-MAR-2002 (first entry)
 DT
 XX *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 30866.
 DE
 XX *Drosophila*, developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS *Drosophila melanogaster*.
 XX WO200171042-A2.
 FN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US092231.
 PE
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 DR WPI; 2001-656860/75.
 DR P-PSDB: ABB68025.
 XX
 PT New, isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 30866; 21bp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX Sequence 9404 BP; 2675 A; 1973 C; 1984 G; 2772 T; 0 other;
 SO
 Query Match 11.0%; Score 35.4; DB 23; Length 9404;
 Best Local Similarity 57.8%; Pred. No. 0.36;
 Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 Oy 43 AGTGAATACCTCTTCTCTCTGCAATCCGCGCGGAGCAAGCAATGGCGCAGTGGCTTA 102
 Db 4027 ATTGGAACCTTTCTTCTCTGCAATCCGCGCGGAGCAATGGCGCAGTGGCTT 4086
 Oy 103 CTTATATACCACTTATATATAGAAAGGAGCGCTCTGGATGCAT 151
 Db 4087 CGTCGCGCAAACTTATATCTAAATACAGCAATGCGCTCGATTAAT 4135
 RESULT 7
 AAX13487/C
 ID AAX13487 standard; DNA; 9797 BP.
 XX
 AC AAX13487;
 XX 19-MAR-1999 (first entry)
 DT
 XX *Enterococcus faecalis* genome contig SEQ ID NO:550.
 DE
 XX *Enterococcus faecalis*; contig; detection; Enterococcal infection;
 KW vaccine; attenuation; computer readable medium; ds.
 XX
 OS *Enterococcus faecalis*.
 XX WO9850555-A2.
 FN
 XX 12-NOV-1998.
 PD
 XX 04-MAY-1998; 98WO-US08985.
 PE
 XX 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Barash SC, Dillon PJ, Kunsch CA;
 DR WPI; 1999-045171/04.
 DR
 XX
 PT New isolated *Enterococcus faecalis* polynucleotides and polypeptides
 PT - used to develop products for the detection of *Enterococcus* and for
 PT use in vaccines for prevention or attenuation of *Enterococcus*
 PT infection.
 XX
 PS Claim 1; Page 1805-1810; 2084pp; English.
 CC A computer readable medium has been developed which has recorded on it
 CC 982 nucleotide sequences isolated from the *Enterococcus faecalis* genome.
 CC AAX12938 to AAX13919 represent these nucleotide sequences which are
 CC primary nucleotide sequences, also known as contigs. The computer-based
 CC system can identify fragments of the *Enterococcus faecalis* genome with
 CC commercial importance. The products can be used to detect the presence
 CC of *Enterococcus faecalis* in samples. They can also be used for
 CC diagnosing *Enterococcal* infection in an animal and monitoring
 CC progression of disease, and for identifying agents which can be used to
 CC modulate the growth or pathogenicity of *Enterococcus faecalis*, or
 CC another related organism, in vivo or in vitro. In particular the
 CC polypeptides encoded by the *Enterococcus faecalis* nucleotide sequences
 CC can be used in vaccines to prevent or attenuate an *Enterococcal*
 CC infection.

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral ischaemia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 XX
 SQ Sequence 1572 BP: 333 A: 378 C: 440 G: 421 T: 0 other;
 Query Match 10.1%; Score 32.4; DB 21; Length 1572;
 Best Local Similarity 49.4%; Pred. No. 1.7;
 Matches 84; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
 OY 86 AATGGCGAGTGGCTTACTTATATACACCAACTATATACAGAAAGCGAGCGTCTGG 145
 DB 71 AAGGGGCGATGAGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 130
 OY 146 ATCGATTGCAAAATCTACCTCCAAACCAACCCAGCTTTGTATCTGCTTACTGATCA 205
 DB 131 AGAAAGGTGACCCCTGATCCCAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190
 OY 206 AGTTGCTGCTGATACGATCTGCTATATGCTCTCTCTCTCTCTCTCTCTCTCTCT 255
 DB 191 TGTCTTTTGTGTAACGTGCTGCTGCAATGTTGCTCTTTGTGCAATATCTGTT 240
 RESULT 10
 AAX00461
 ID AAX00461 standard; cDNA; 4942 BP.
 AC AAX00461;
 DT 21-MAY-1999 (first entry)
 XX
 DE Human type VI adenyl cyclase cDNA.
 KW Adenyl cyclase type VI; human; hAC6; therapy; diagnosis; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 145 3651
 FT /*tag= a
 XX
 XX W09901547-A1.
 XX
 XX 14-JAN-1999.
 XX
 XX 01-JUL-1998; 98MO-US13694.
 XX
 XX 01-JUL-1997; 97US-0886550.
 XX
 XX 01-JUL-1997; 97US-0070904.
 XX
 XX (COR-) COR THERAPEUTICS INC.
 XX
 XX Tomlinson JA.
 XX
 XX WPI: 1999-106049/09.
 XX
 XX P-PSDB: AAM30599.
 XX
 XX
 PT Newly isolated and purified human type VI adenyl cyclase (hAC6)
 PT polypeptide - useful for identifying potential therapeutic agents
 PT that modulate hAC6 activity, and for the diagnosis of
 PT hAC6-associated diseases and disorders
 PS
 PS Claim 1; Fig 1A-1; 42pp; English.
 CC
 CC This DNA sequence encodes human type VI adenyl cyclase (hAC6, see
 CC AM33599) that is expressed mainly in the heart and brain. hAC6 has
 CC a similar putative structure to other adenyl cyclase isoforms
 CC but, like type V, is distinguishable in that it has a larger

CC N-terminus and a relatively shorter C-terminus as it lacks the C2b
 CC region. hAC6 cDNA was initially isolated from a human heart cDNA
 CC library using an adenyl cyclase PCR fragment as probe. It was
 CC used to design primers that were used in a PCR-based RACE to obtain
 CC the full-length cDNA sequence. The invention relates to the hAC6
 CC gene, methods for the recombinant production of purified hAC6 and
 CC the proteins made by these methods, antibodies against hAC6,
 CC vectors, probes and host cells (especially HEK-293) transformed by
 CC genes encoding polypeptides having hAC6 activity, along with
 CC diagnostic and therapeutic uses for these various reagents. hAC6
 CC can be used as a tool to screen for agonists and antagonists that
 CC stimulate/inhibit hAC6. Such compounds have therapeutic utility
 CC in treating diseases caused by aberrant activity of this enzyme,
 CC and diseases whose symptoms can be ameliorated by stimulating or
 CC inhibiting the activity of hAC6.
 XX
 SQ Sequence 4942 BP: 953 A: 1404 C: 1512 G: 1073 T: 0 other;
 Query Match 10.1%; Score 32.4; DB 20; Length 4942;
 Best Local Similarity 49.4%; Pred. No. 2.9;
 Matches 84; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
 OY 86 AATGGCGAGTGGCTTACTTATATACACCAACTATATACCTAATAAAGCGAGCTCTGG 145
 DB 4559 AAGGGGCGATGAGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4618
 OY 146 ATCGATTGCAAAATCTACCTCCAAACCAACCCAGCTTTGTATCTGCTTACTGATCA 205
 DB 4619 AGAAAGGTGACCCCTGATCCCAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4678
 OY 206 AGTTGCTGCTGATACGATCTGCTATATGCTCTCTCTCTCTCTCTCTCTCTCTCT 255
 DB 4679 TGTCTTTTGTGTAACGTGCTGCTGCAATGTTGCTCTTTGTGCAATATCTGTT 4728
 RESULT 11
 AAD33336/C
 ID AAD33336 standard; DNA; 3182 BP.
 AC AAD33336;
 DT 01-JUL-2002 (first entry)
 XX
 DE A. thaliana SAG12 promoter/1PT/NOS-ter construct (alternative version).
 KW Senescence associated gene; SAG-12; transgenic plant; transgenic;
 KW senescence-specific promoter; senescence characteristic; chimeric;
 KW IPT; isopenentenyl transferase; NOS-ter; ds.
 XX
 OS Chimeric - Arabidopsis thaliana.
 OS Chimeric - Unidentified.
 XX
 XX US6359197-B1.
 XX
 XX 19-MAR-2002.
 XX
 XX 17-NOV-1997; 97US-0971395.
 XX
 XX 29-MAR-1995; 95US-0413135.
 XX
 XX (WISC) WISCONSIN ALUMNI RES FOUND.
 XX
 XX Amasino RM, Gan S, Noh Y;
 XX
 XX WPI: 2002-291012/33.
 XX
 XX
 PT Novel senescence associated promoter sequence connected to a
 PT protein-coding DNA sequence useful for the creation of transgenic
 PT plants with altered senescence characteristics -
 PS
 PS Claim 1; Column 15-20; 21pp; English.
 CC
 CC The invention relates to a genetic construct comprising senescence


```

PR 15-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144085;
PR 19-JUL-1999; 99US-0144086;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144632;
PR 20-JUL-1999; 99US-0144884;
PR 21-JUL-1999; 99US-0144814;
PR 21-JUL-1999; 99US-0145086;
PR 21-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145085;
PR 22-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146388;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 10-AUG-1999; 99US-0147935;
PR 11-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 23-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 30-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;

```

```

PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158363;
PR 13-OCT-1999; 99US-0158393;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159337;
PR 14-OCT-1999; 99US-0159637;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161922;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

```

```

Query Match          9.9%; Score 31.8; DB 21; Length 1894;
Best Local Similarity 67.2%; Pred. No. 3;
Matches 45; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

```

QY      14  GCGGTCACATGACCGTAACAGTGCAGTGCATCTTCTCTCGAATCCGT 73
          |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db       133  GCGGTCACATGACCGTAACAGTGCAGTGCATCTTCTCTCGAATCCGT 74

QY      74  GCGGTCG 80
          |||  |||
Db       73  GCGGCGG 67

```

```

RESULT 15
AAK59121/c
ID AAK59121 standard; cDNA; 1947 BP.

```

```

XX      AAK59121,
AC
XX
DT      31-AUG-1999 (first entry)

```

```

DE      Maize UDP-glucose dehydrogenase cDNA clone Zmudpgh2.

```

```

XX      UDP-glucose dehydrogenase; UDPgdh; maize; corn; transgenic plant;
KM      starch; disease resistance; pest resistance, ss.
XX

```

```

OS      Zea mays.

```

```

XX      Key          Location/Qualifiers
FT      CDS          174..1616
FT      FT           /*lag-a
FT      FT           /EC_number="1.1.1.22"

```

```

XX      MO9929875-A2.

```

```

XX      17-JUN-1999.

```

```

XX      09-DEC-1998; 98MO-US26190.

```

```

XX      10-DEC-1997; 97US-0987367.

```

XX (PION) PIONEER HI-BRED INT INC.

PA Bruce WR, Dhuga KS, Lu G, Nichols SE, Saunders CA;
 PI Sims LE, Singletary GW, Witcher D, Zhong G;

XX WPI: 1999-385615/32.
 DR P-PSDE; AAY06307.

XX New maize UDP-glucose dehydrogenase useful for increasing the
 PT biosynthesis of gums in seeds of transgenic plants

PS Claim 16; Fig 3; 82pp; English.

XX This is the nucleotide sequence of maize UDP-glucose dehydrogenase
 CC (UDPgdH) (NC 1.1.1.22) cDNA clone Zmudpgh2, a polymorphic variant
 CC of clone Zmudpgh1 (see AAX59120). The 2 clones were isolated by PCR
 CC amplification of maize cDNA using M13 primers and identified as
 CC UDPgdH sequences using a BLAST program. The invention provides
 CC maize UDPgdH, its variants and mutants, especially a mutant in
 CC which the Cys-272 residue of UDPgdH is replaced by another amino
 CC acid residue. UDPgdH (see AAY06306) catalyses a 2-step reaction in
 CC which UDP-D-glucose is converted to UDP-6-aldehydo-D-glucose, which
 CC is in turn converted to UDP-D-glucuronic acid. With the mutant
 CC enzymes, UDP-6-dehydro-D-glucose is the end product. UDPgdH
 CC nucleic acids can be used to produce transgenic plants having
 CC altered quality or quantity of starch. Expression of the mutant
 CC enzyme provides aldehydic starch. A vector in which a promoter is
 CC operatively linked to a UDPgdH nucleic acid is used in claimed
 CC methods for increasing the stalk or stem length of plants, for
 CC increasing insect and pathogen resistance, for increasing root
 CC strength, and for increasing resistance to root worm in plants.
 CC A vector in which a promoter is linked to a gene that inhibits
 CC UDPgdH enzyme activity (e.g. an antisense or ribozyme sequence) is
 CC used in methods for increasing the extractability of starch in corn
 CC wet milling, and for increasing the nutritional value of plants
 CC or plant seeds.

CC Sequence 1947 BP; 420 A; 566 C; 539 G; 422 T; 0 other;

SO Query Match

Best Local Similarity 9.9%; Score 31.8; DB 20; Length 1947;

Matches 45; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 14 GCGGTCGACTGAACCTTAACAGTGAAGATGATCTCTCTCTGCAATCGCT 73
 DB 113 GCGGTCGACTGAACCTTAACAGTGAAGATGATCTCTCTCTGCAATCGCT 73
 OY 74 GCGGTG 80
 DB 53 GCGGCGG 47

Search completed: August 18, 2003, 13:23:09
 Job time : 216 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 12:28:30 ; Search time 2026 Seconds
(without alignments)
6501.925 Million cell updates/sec

Title: US-10-058-566-5
Perfect score: 322
Sequence: 1 gaattctgcgtcgcgcgtc.....agtgtctagctgagaacatg 322

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: GenEmbl:*
2: gb_pa:*
3: gb_hg:*
4: gb_in:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgc_hum:*
31: em_hgc_inv:*
32: em_hgc_other:*
33: em_hgc_mus:*
34: em_hgc_pln:*
35: em_hgc_rdt:*
36: em_hgc_mam:*
37: em_hgc_vrt:*
38: em_sy:*
39: em_hgc_hum:*
40: em_hgc_inv:*
41: em_hgc_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 322 | 100.0 | 322 | 6 | AX512960 Sequence |
| 2 | 322 | 100.0 | 322 | 8 | AF366294 Zea mays |
| 3 | 322 | 100.0 | 2541 | 6 | AX512958 Sequence |
| 4 | 187 | 58.1 | 187 | 6 | AX512961 Sequence |
| 5 | 39.4 | 12.2 | 145550 | 2 | AC099403 Oryza sat |
| 6 | 39.4 | 12.2 | 173729 | 2 | AC108757 Oryza sat |
| 7 | 38.8 | 12.0 | 202765 | 2 | BX465192 Danio rer |
| 8 | 38.8 | 12.0 | 204254 | 2 | BX465854 Danio rer |
| 9 | 37.8 | 11.7 | 179373 | 10 | AL591606 Mouse DNA |
| 10 | 36.8 | 11.4 | 2000 | 6 | AX655393 Sequence |
| 11 | 36.2 | 11.2 | 192623 | 2 | AC131032 Mus muscu |
| 12 | 36 | 11.2 | 112292 | 9 | AC084297 Homo sapi |
| 13 | 36 | 11.2 | 153774 | 2 | AC021009 Homo sapi |
| 14 | 35.6 | 11.1 | 132635 | 2 | AC007577 Drosophila |
| 15 | 35.6 | 11.1 | 156805 | 2 | AC020006 Drosophila |
| 16 | 35.6 | 11.1 | 199016 | 2 | AC008204 Drosophila |
| 17 | 35.6 | 11.1 | 206741 | 3 | AE003747 Drosophila |
| 18 | 35.6 | 11.1 | 250495 | 10 | AL844873 Mouse DNA |
| 19 | 35.4 | 11.0 | 5933 | 3 | AF246991 Drosophila |
| 20 | 35.4 | 11.0 | 6352 | 3 | AY119520 Drosophila |
| 21 | 35.4 | 11.0 | 38193 | 2 | AC017671 Drosophila |
| 22 | 35.4 | 11.0 | 167977 | 3 | AC010010 Drosophila |
| 23 | 35.4 | 11.0 | 168250 | 3 | AC093121 Drosophila |
| 24 | 35.4 | 11.0 | 300610 | 3 | AE003467 Drosophila |
| 25 | 35.2 | 10.9 | 180189 | 2 | AC018864 Homo sapi |
| 26 | 35.2 | 10.9 | 182855 | 9 | AC009292 Homo sapi |
| 27 | 35.2 | 10.9 | 186304 | 9 | AC090136 Homo sapi |
| 28 | 35.2 | 10.9 | 237705 | 9 | AF215845 Homo sapi |
| 29 | 35 | 10.9 | 43886 | 2 | AC021668 Homo sapi |
| 30 | 34.8 | 10.8 | 171708 | 9 | AC067956 Homo sapi |
| 31 | 34.4 | 10.7 | 24725 | 3 | CEK0108 Caenorhabdi |
| 32 | 34.4 | 10.7 | 141084 | 10 | AC127971 Rattus no |
| 33 | 34.4 | 10.7 | 166167 | 10 | AL611936 Rattus no |
| 34 | 34.4 | 10.7 | 180291 | 10 | AC129146 Rattus no |
| 35 | 34.4 | 10.7 | 248625 | 2 | AC106189 Rattus no |
| 36 | 34.4 | 10.7 | 248675 | 2 | AL844206 Mus muscu |
| 37 | 34.4 | 10.7 | 310448 | 2 | AC131354 Rattus no |
| 38 | 34.2 | 10.6 | 171930 | 2 | AC118687 Mus muscu |
| 39 | 34.2 | 10.6 | 208765 | 2 | BX004985 Mus muscu |
| 40 | 34 | 10.6 | 59841 | 9 | HS180M12 Human DNA |
| 41 | 34 | 10.6 | 232610 | 2 | AC123157 Rattus no |
| 42 | 34 | 10.6 | 232631 | 2 | AC110413 Rattus no |
| 43 | 34 | 10.6 | 247076 | 2 | AC119394 Rattus no |
| 44 | 33.8 | 10.5 | 113391 | 2 | AC119985 Mus muscu |
| 45 | 33.8 | 10.5 | 161506 | 9 | AL450105 Human DNA |

ALIGNMENTS

RESULT 1
AX512960
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AX512960 322 bp DNA linear PAT 03-OCT-2002
Sequence 5 from Patent WO02063021.
AX512960.1 GI:23504095
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
Albertsen, M., Fox, T., Hufman, G., and Trimel, M.
Nucleotide sequence mediating male fertility and method of using

Pred. No. is the number of results predicted by chance to have a

same
JOURNAL Patent: WO 02063021-A 5 15-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
SOURCE
1. .322
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
BASE COUNT 79 a 83 c 69 g 91 t
ORIGIN

Query Match 100.0%; Score 322; DB 6; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.1e-93;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTCTCT 60
DB 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTCTCT 60
QY 61 CTCTGCAATTCCTGCTGAGAGTAAATGAGTGGTCTCTACTTATCAACCACTTAT 120
DB 61 CTCTGCAATTCCTGCTGAGAGTAAATGAGTGGTCTCTACTTATCAACCACTTAT 120
QY 121 CACCTAGAAAAGCAGCGCTCTGATGATGTAATCTACTCTCCACCAACCCAGCTT 180
DB 121 CACCTAGAAAAGCAGCGCTCTGATGATGTAATCTACTCTCCACCAACCCAGCTT 180
QY 131 TGTATCTGCTTACTGCTATGATGACCAAGTGTGCTGATGATGCTGATTTCTTT 240
DB 131 TGTATCTGCTTACTGCTATGATGACCAAGTGTGCTGATGATGCTGATTTCTTT 240
QY 241 CTCTCTAGAAATGTTCTGCGGATGCTTTATAGAGAGGTTGGTCAATGATCTTG 300
DB 241 CTCTCTAGAAATGTTCTGCGGATGCTTTATAGAGAGGTTGGTCAATGATCTTG 300
QY 301 CCAAGTGTCTAGTGAACATG 322
DB 301 CCAAGTGTCTAGTGAACATG 322

KRSJUT 2
AF366294
LACUS
DEFINITION Zea mays dihydro-flavanoid reductase-like protein (ms*-bs7) gene,
promoter sequence.
ACCESSION AF366294
VERSION AF366294.1 GI:14030552
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
1 (bases 1 to 322)
AUTHORS Fox,T.W., Trimnell,M.P. and Albertsen,M.C.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7100 N.W. 62nd Ave. P O Box 1004, Johnston, IA
50131-1004, USA

FEATURES
SOURCE
1. .322
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
1. .>322
/gene="ms*-bs7"
/note="dihydro-flavanoid reductase-like protein; anther
specific"
1. .322
/gene="ms*-bs7"
misc_feature
165..174
/gene="ms*-bs7"
/note="Myb binding domain"

TATA_signal 269..272
/gene="ms*-bs7"
BASE COUNT 79 a 83 c 69 g 91 t
ORIGIN

Query Match 100.0%; Score 322; DB 8; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.1e-93;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTCTCT 60
DB 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTCTCT 60
QY 61 CTCTGCAATTCCTGCTGAGAGTAAATGAGTGGTCTCTACTTATCAACCACTTAT 120
DB 61 CTCTGCAATTCCTGCTGAGAGTAAATGAGTGGTCTCTACTTATCAACCACTTAT 120
QY 121 CACCTAGAAAAGCAGCGCTCTGATGATGTAATCTACTCTCCACCAACCCAGCTT 180
DB 121 CACCTAGAAAAGCAGCGCTCTGATGATGTAATCTACTCTCCACCAACCCAGCTT 180
QY 181 TGTATCTGCTTACTGCTATGATGACCAAGTGTGCTGATGATGCTGATTTCTTT 240
DB 181 TGTATCTGCTTACTGCTATGATGACCAAGTGTGCTGATGATGCTGATTTCTTT 240
QY 241 CTCTCTAGAAATGTTCTGCGGATGCTTTATAGAGAGGTTGGTCAATGATCTTG 300
DB 241 CTCTCTAGAAATGTTCTGCGGATGCTTTATAGAGAGGTTGGTCAATGATCTTG 300
QY 301 CCAAGTGTCTAGTGAACATG 322
DB 301 CCAAGTGTCTAGTGAACATG 322

RESULT 3
AX512958
LOCUS AX512958 2541 bp DNA linear PAT 03-OCT-2002
DEFINITION Sequence 3 from Patent W002063021.
ACCESSION AX512958
VERSION AX512958.1 GI:23504094
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS Albertsen,M., Fox,T., Huffman,G. and Trimnell,M.
TITLE Nucleotide sequence mediating male fertility and method of using
same
JOURNAL Patent: WO 02063021-A 3 15-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
SOURCE
1. .2541
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
BASE COUNT 577 a 693 c 631 g 640 t
ORIGIN

Query Match 100.0%; Score 322; DB 6; Length 2541.
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTCTCT 60
DB 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTCTCT 60
QY 61 CTCTGCAATTCCTGCTGAGAGTAAATGAGTGGTCTCTACTTATCAACCACTTAT 120
DB 61 CTCTGCAATTCCTGCTGAGAGTAAATGAGTGGTCTCTACTTATCAACCACTTAT 120
QY 121 CACCTAGAAAAGCAGCGCTCTGATGATGTAATCTACTCTCCACCAACCCAGCTT 180

```

Db      121 CACGTAGAAAGGAGCGCTCGATCGATTGCAATCTACTCCACCAACCAAGCTT 180
QY      181 TGTATCTGCTTACTGCTGATGACCAAGTTGCTCTATAGATGTCGATTTGCTTT 240
Db      181 TGTATCTGCTTACTGCTGATGACCAAGTTGCTCTATAGATGTCGATTTGCTTT 240
QY      241 CTCTCTAGAAAGTTCTGCGCGATGCTTATAGAGAAGTTGGTCACATCGATCTGTG 300
Db      241 CTCTCTAGAAAGTTCTGCGCGATGCTTATAGAGAAGTTGGTCACATCGATCTGTG 300
QY      301 CCAGTCTAGATGTTCTGCGATGCTTATAGAGAAGTTGGTCACATCGATCTGTG 300
Db      301 CCAGTCTAGATGTTCTGCGATGCTTATAGAGAAGTTGGTCACATCGATCTGTG 300

RESULT 4
AX512961 187 bp DNA linear PAT 03-OCT-2002
LOCUS     Sequence 6 from Patent WO02063021.
DEFINITION
ACCESSION AX512961
VERSION    AX512961.1 GI:23504096
KEYWORDS
SOURCE     Zea mays
ORGANISM   Zea mays
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
           clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS   Albertsen, M., Fox, T., Hoffman, G. and Trimmell, M.
TITLE      Nucleotide sequence mediating male fertility and method of using
           same
JOURNAL    Patent: WO 02063021-A 6 15-AUG-2002;
           PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
   source          1..187
                   /organism="Zea mays"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:4577"
BASE COUNT  42 a 46 c 40 g 59 t
ORIGIN
Query Match      58.1%; Score 187; DB 6; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.2e-49;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      136 CGGCTCTGATGATGCAATCTACCTCAACCAACCAAGCTTGTATGCTTACTG 195
Db      1 CGGCTCTGATGATGCAATCTACCTCAACCAACCAAGCTTGTATGCTTACTG 60
QY      196 TGATCACCAGTTGTGCTGATAGATGTCGATATTGCTCTTCTCTAGAAATGT 255
Db      61 TGATCACCAGTTGTGCTGATAGATGTCGATATTGCTCTTCTCTAGAAATGT 120
QY      256 CCGCGCGATGCTTATAGAGAAGTTGGTCACATCGATCTGCGAGTGTCTAGCTGA 315
Db      121 CCGCGCGATGCTTATAGAGAAGTTGGTCACATCGATCTGCGAGTGTCTAGCTGA 180
QY      316 GAACATG 322
Db      181 GAACATG 187

```

```

REFERENCE
AUTHORS   Yun, D.-W., Hahn, J.-H., Yoon, U.-H., Lee, J.-S., Lee, M.-C., Eun, M.Y.
           and Kim, H.-I.
TITLE      Oryza sativa BAC OSJNBa005C18 genomic sequence
JOURNAL    Unpublished
REFERENCE 2 (bases 1 to 145550)
AUTHORS   Hahn, J.-H. and Kim, H.-I.
TITLE      Direct Submission
JOURNAL    Submitted (14-NOV-2001) Rice Genome Sequencing Project, National
           Institute of Agricultural Science and Technology (NIAS), RDA, 249
           Seodun-dong, Suwon 441-707, Korea (E-mail: jhahn@nias.go.kr,
           Tel: 82-31-290-0309, Fax: 82-31-290-0308)
COMMENT   * NOTE: This is a 'working draft' sequence. It currently
           * consists of 5 contigs. Gaps between the contigs
           * are represented as runs of N. The order of the pieces
           * is believed to be correct as given, however the sizes
           * of the gaps between them are based on estimates that have
           * provided by the submitter.
           * This sequence will be replaced
           * by the finished sequence as soon as it is available and
           * the accession number will be preserved.
           * 1 86000: contig of 86000 bp in length
           * 86001 86100: gap of unknown length
           * 86101 111963: contig of 25863 bp in length
           * 111964 112063: gap of unknown length
           * 112064 116883: contig of 7620 bp in length
           * 116884 119783: gap of unknown length
           * 119784 144017: contig of 24234 bp in length
           * 144018 144117: gap of unknown length
           * 144118 145550: contig of 1433 bp in length.
           *

FEATURES
   source          1..145550
                   /organism="Oryza sativa (japonica cultivar-group)"
                   /mol_type="genomic DNA"
                   /variety="Nipponbare"
                   /db_xref="taxon:39947"
                   /chromosome="9"
                   /clone="OSJNBa005C18"
BASE COUNT  38466 a 33520 c 33408 g 39755 t 401 others
ORIGIN
Query Match      12.2%; Score 39.4; DB 2; Length 145550;
Best Local Similarity 60.6%; Pred. No. 0.27;
Matches 83; Conservative 0; Mismatches 51; Indels 3; Gaps 1;

QY      156 AATCTACCTCAACCAACCAAGCTTGTATGCTTACTGATCACCACCAAGTTGCTG 215
Db      66598 AATCTACCTCAACCAACCAAGCTTGTATGCTTACTGATCACCACCAAGTTGCTG 215
QY      216 ATAGCATGTGATATGCTCTTCTCTCTAGAAAGTTGCTGCGATGCTTATAGAGA 275
Db      66655 TGATGTACAAATTTCTGCTCTCTCTCTAGAAAGTTGCTGCGATGCTTATAGAGA 66714
QY      276 GAAGTTGGTCACATG 232
Db      66715 GAAGTTGGTCACATG 66731

RESULT 6
AC108757/c 173729 bp DNA linear HTG 31-JAN-2002
LOCUS     Oryza sativa (japonica cultivar-group) chromosome 9 clone
DEFINITION
ACCESSION AC108757
VERSION    AC108757.1 GI:18449962
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```


QY 177 GCTTTGATTCCTTACTGCTGATCACAAGTTGCTGATACGATGTCGATTATTCCT 236
DB 55054 GCACCTAGCTTTTCATGTCGTCCATCCTGTACAAATATATATGACTTATTCGA 55113
QY 237 CT 238
DB 55114 CT 55115

RESULT 8
BX465854 204254 bp DNA linear HTG 03-JUN-2003
DEFINITION Danio rerio clone DKRY-208M10, *** SEQUENCING IN PROGRESS ***.
ACCESSION BX465854 GI:31408148
VERSION BX465854.4
KEYWORDS HTG: HTGS_PHASE2: HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 204254)
McLaren, S.
Direct Submission
Submitted (01-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jun 4, 2003 this sequence version replaced gi:30524774.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK208M10
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 204134 bases at least Q40
Consensus quality: 204198 bases at least Q30
Consensus quality: 204237 bases at least Q20
Insert size: 204254; sum-of-contigs
Quality coverage: 210284; 2.8% error; agarose-fp
Quality coverage: 8.60x in Q20 bases; sum-of-contigs quality
coverage: 8.35x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 204254. contig of 204254 bp in length.
Location/Qualifiers
1. 204254
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKRY-208M10"
/clone_1lb="DanioKey"
1. 204254
/note="assembly_fragment:00247
clone end:SP6
vector side:left"
BASE COUNT 67247 a 37099 c 36096 g 63812 t
ORIGIN

Query Match 12.0%; Score 38.8; DB 2; Length 204254;
Best Local Similarity 57.4%; Pred. No. 0.43;
Matches 70; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 117 TTATCAGTAAAGCGAGCGCTCGATGATTCGAATTCACGCCAACCAACCA 176
DB 116241 TTAGACCAACACACACACAGGTTAGCGCCCTGCTTACTTAACCAATTC 116300
QY 177 GCTTTGATTCCTTACTGCTGATCACAAGTTGCTGATACGATGTCGATTATTCCT 236
DB 116301 GCACCTAGCTTTTCATGTCGTCCATCCTGTACAAATATATATGACTTATTCGA 116360
QY 237 CT 238
DB 116361 CT 116362

RESULT 9
AL591606 179373 bp DNA linear ROD 05-APR-2002
LOCUS Mouse DNA sequence from clone RP23-272014 on chromosome 2, complete
DEFINITION sequence.
ACCESSION AL591606
VERSION AL591606.4 GI:20068478
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Almeida, J.
Direct Submission
Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:19571573.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-272014 is
from the RPI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.
Location/Qualifiers
1. 179373
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-272014"
/clone_1lb="RPI-23"
BASE COUNT 49498 a 37562 c 39607 g 52706 t
ORIGIN

Query Match 11.7%; Score 37.8; DB 10; Length 179373;
Best Local Similarity 54.7%; Pred. No. 0.91;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 117 TTATCAGTAAAGCGAGCGCTCGATGATTCGAATTCACGCCAACCAACCA 176
DB 37438 TTCTAACCTTAAGAAATGAGCTGCGCTTGTCGAATTCACCACTACATCCCAAGCA 37379

QY 177 GCTTGTATCTGCTTACGTCGATCACCAGATTGCTGATACGATGCTGATATTGCT 236
 DB 37378 TAGGTGTTCTTCCATGCTGCTGATGATGGCCCATCATCATGAGCAAGATAGAC 37319
 QY 237 CTTTCTTCTCTAGATG 253
 DB 37318 CAGCCTTCACAGAAATG 37302
 RESULT 10
 AK655393/1
 LOCUS AK655393 2000 bp DNA linear PAT 22-MAR-2003
 DEFINITION Sequence 5263 from Patent WO03000898.
 ACCESSION AK655393
 VERSION AK655393.1 GI:29158207
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa
 Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Euphorbiaceae; Oryzae; Oryza.
 REFERENCE
 1 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T., and Zou, G.,
 Plant genes involved in defense against pathogens
 Patent: WO 03000898-A 5263 03-JAN-2003;
 Syngenta Participations AG (CH)
 FEATURES
 source
 1. 2000
 /organism="Oryza sativa"
 /mol_type="genomic DNA"
 /db_xref="taxon:4530"
 BASE COUNT 336 a 265 c 284 g 363 t 752 others
 ORIGIN
 Query Match 11.4%; Score 36.8; DB 6; Length 2000;
 Best local Similarity 9.5%; Pred. No. 1.4;
 Matches 23; Conservative 121; Mismatches 98; Indels 0; Gaps 0;
 QY 77 GTGAGCAAAATGCGGAGTGGCTTACTTATCAGACCACTTATCAGTAAAGGAC 136
 DB 568 KKKMKRRMYMMYMKCTWRRCMCYRMQCYMTTTSRSMYTGKRYKARYTSKRMYMYKYR 509
 QY 137 GCGTCTGATGATTCATTCATCTACTCTCAACCAACCCAGCTTGTATGCTTACTGT 196
 DB 508 KTCYTYTYGMYKMCYSMMRYGCKACKCCYACMCWKAAYGMMYWKYRKYSKMRMSTKY 449
 QY 197 GATCACCAGATTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 256
 DB 448 MMSMYKKKSRMYGAKGCGKMMWTCSYGYMKWTYMSYKRYKMYKMYKMYKMYKMYK 389
 QY 257 CTGCGGATGCTTATTAAGAGAGAGTGTGATGATGATGATGATGATGATGATGATG 316
 DB 338 YMMYYSAYSSMMYTYYYAKYMYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYK 329
 QY 317 AA 318
 DB 328 MM 327
 RESULT 11
 AC131032 192623 bp DNA linear HTG 22-MAR-2003
 LOCUS AC131032
 DEFINITION Mus musculus clone RP23-347J22, WORKING DRAFT SEQUENCE, 7 unordered
 pieces.
 ACCESSION AC131032
 VERSION AC131032.4 GI:29150473
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

1 (bases 1 to 192623)
 Birren, B., Nusbaum, C., and Lander, E.
 Mus musculus, clone RP23-347J22
 Unpublished
 2 (bases 1 to 192623)
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choquel, Y., Collumore, A.,
 Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, S., Dodge, S.,
 Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, P., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, C., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Notly, C., Norman, C. H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rice, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 192623)
 Birren, B., Nusbaum, C., Lander, E., Abouneil, A., Allen, N.,
 Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choquel, Y.,
 Collumore, A., Cook, A., Cooke, P., Corum, B., Deatellano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S.,
 Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hatz, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, P., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, C., Liu, A., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rice, C., Rogov, P.,
 Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (22-MAR-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 22, 2003 this sequence version replaced g1:28269650.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L26960
 Center clone name: 347.J.22
 Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 189879 bases at least Q40
 Consensus quality: 190483 bases at least Q30
 Consensus quality: 191015 bases at least Q20
 Insert size: 188000; agarose-fp
 Insert size: 192023; sum-of-ctrls
 Quality coverage: 16.0 in Q20 bases; agarose-fp
 Quality coverage: 15.7 in Q20 bases; sum-of-ctrls

the clone sequenced to the right is RP11-279115. Actual start of this clone is at base position 125378 of RP11-545M1; actual end is at base position 112292 of RP11-44H9.

Data from AC0105342 was used to finish the clone, AC084297.

FEATURES

```

source
    location/Qualifiers
    1..112292
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="2"
    /map="2"
    /clone="RP11-44H9"
    /clone_id="RP11-11"
    1490..1569
    /rpt_family="MIR"
    1594..1785
    /rpt_family="MERL_type"
    2224..2403
    /rpt_family="MaLR"
    2413..3725
    /rpt_family="MaLR"
    3735..4157
    /rpt_family="MaLR"
    5426..5571
    /rpt_family="MIR"
    5580..5949
    /rpt_family="ERV1"
    7322..7788
    /rpt_family="ERV1"
    8080..8284
    /rpt_family="MIR"
    8330..8419
    /rpt_family="L1"
    8520..8585
    /rpt_family="L1"
    8593..8637
    /rpt_family="AT_rich"
    11650..11700
    /rpt_family="T-rich"
    16745..16842
    /rpt_family="MaLR"
    19122..19483
    /rpt_family="MaLR"
    19484..20992
    /rpt_family="MaLR"
    21000..21539
    /rpt_family="ERV1"
    21788..22338
    /rpt_family="L1"
    22461..23642
    /rpt_family="L1"
    23042..23066
    /rpt_family="AT_rich"
    23124..23149
    /rpt_family="AT_rich"
    23282..23303
    /rpt_family="AT_rich"
    23643..23924
    /rpt_family="Alu"
    23925..24550
    /rpt_family="L1"
    24635..24657
    /rpt_family="AT_rich"
    24669..24750
    /rpt_family="(TA)n"
    24751..27144
    /rpt_family="L1"
    25921..25949
    /rpt_family="AT_rich"
    26637..26665
    /rpt_family="(A)n"
    26925..26988
    repeat_region

```

```

repeat_region      /rpt_family="(TA)n"
27185..27959      /rpt_family="L1"
27958..28112      /rpt_family="L1"
30417..30577      /rpt_family="L1"
30624..30757      /rpt_family="L1"
30810..31181      /rpt_family="MaLR"
31352..32138      /rpt_family="L1"
32212..32476      /rpt_family="L1"
32477..33286      /rpt_family="ERV1"
33314..33996      /rpt_family="L1"
33995..35252      /rpt_family="L1"
34666..34688      /rpt_family="L1"
35253..35552      /rpt_family="AT_rich"
35553..36104      /rpt_family="Alu"
35836..35898      /rpt_family="L1"
36563..36664      /rpt_family="A-rich"
36894..36982      /rpt_family="MIR"
37044..37602      /rpt_family="MIR"
40321..40413      /rpt_family="MER2_type"
40437..40578      /rpt_family="(TTCC)n"
40808..40935      /rpt_family="L2"
41338..41583      /rpt_family="AchoBo"
42197..42997      /rpt_family="MIR"
42309..42331      /rpt_family="L2"
43050..43346      /rpt_family="AT_rich"
43785..43819      /rpt_family="Alu"
repeat_region

```

Query Match 11.2% Score 36; DB 9; Length 112292;
 Best Local Similarity 60.0%; Pred. No. 3.4;
 Matches 60; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```

QY 126 AGAAACGACGCGCTCGATGATGCAATATCTACCTCCACCAACCCAGCTTGTAT 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 87439 ATATAGTACTCTCTCTCTCTCTGTTGAAATTACCTCATGAAATTAAGCCATGAT 87498
QY 186 CTGCTTACTGTCATCAACCAAGTGTGCGATACGATG 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 87499 CCCCTTATCTTAGAGAGTACTGTGTCATGAGGTGAG 87538

```

RESULT 13
 AC021009/c 153774 bp DNA linear HTG 10-SEP-2000
 LOCUS Homo sapiens clone RP11-279115, WORKING DRAFT SEQUENCE, 10
 DEFINITION unorderd pieces.
 ACCESSION AC021009
 VERSION AC021009.4 GI:10047819
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)

TITLE
JOURNAL
REFERENCE
AUTHORS

Moshirell, A.R., Moshirell, M., Nixon, K., Pacle, J.M., Park, S.,
Pleiffer, B., Poon, L., Sequelra, A., Sethi, H., Solr, E.,
Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and
Rubin, G.M.
Sequencing of *Drosophila melanogaster*
2 (bases 1 to 132635)
Celniker, S.E., Agbayani, A., Arcalena, T.T., Baxter, E., Blazer, F.G.,
Boyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Katta, K., Kearney, L.,
Kam, E., Lee, B., Lewis, S., Li, P., Lomoran, M.A., Mazda, P.,
Moshirell, A.R., Moshirell, M., Nixon, K., Pacle, J.M., Park, S.,
Pleiffer, B., Poon, L., Sequelra, A., Sethi, H., Solr, E.,
Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and
Rubin, G.M.
Submitted (18-MAY-1999) *Drosophila* Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Aug 2, 1999 this sequence version replaced q1:5629990.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to hdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 87 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 807: contig of 807 bp in length
* 808 887: gap of unknown length
* 1606 1605: contig of 718 bp in length
* 1686 2325: gap of unknown length
* 2326 2405: gap of unknown length
* 2406 3474: contig of 1069 bp in length
* 3475 3554: gap of unknown length
* 3555 4424: contig of 870 bp in length
* 4425 4504: gap of unknown length
* 4505 5429: contig of 925 bp in length
* 5430 5509: gap of unknown length
* 5510 6622: contig of 1113 bp in length
* 6623 6702: gap of unknown length
* 6703 7819: contig of 1117 bp in length
* 7820 7899: gap of unknown length
* 7900 9031: contig of 1132 bp in length
* 9032 9111: gap of unknown length
* 9112 9994: contig of 883 bp in length
* 9995 10074: gap of unknown length
* 10075 10863: contig of 789 bp in length
* 10864 10943: gap of unknown length
* 10944 11546: contig of 603 bp in length
* 11547 11626: gap of unknown length
* 11627 12586: contig of 960 bp in length
* 12587 12666: gap of unknown length
* 12667 13832: contig of 1166 bp in length
* 13833 13912: gap of unknown length
* 13913 14559: contig of 647 bp in length
* 14560 14639: gap of unknown length
* 14640 15453: contig of 814 bp in length
* 15454 15533: gap of unknown length
* 15534 16708: contig of 1176 bp in length
* 16710 16789: gap of unknown length
* 16790 17413: contig of 624 bp in length
* 17414 17493: gap of unknown length
* 17494 18143: contig of 650 bp in length
* 18144 18223: gap of unknown length
* 18224 19349: contig of 1106 bp in length
* 19350 19409: gap of unknown length

19410 20351: contig of 942 bp in length
* 20352 20431: gap of unknown length
* 20432 21257: contig of 826 bp in length
* 21258 21337: gap of unknown length
* 21338 22272: contig of 935 bp in length
* 22273 22352: gap of unknown length
* 22353 23492: contig of 1140 bp in length
* 23493 23572: gap of unknown length
* 23573 24323: contig of 750 bp in length
* 24324 24403: gap of unknown length
* 24403 25741: contig of 1339 bp in length
* 25742 25821: gap of unknown length
* 25822 27015: contig of 1194 bp in length
* 27016 27095: gap of unknown length
* 27096 27979: contig of 884 bp in length
* 27980 28059: gap of unknown length
* 28060 29090: contig of 1031 bp in length
* 29091 29170: gap of unknown length
* 29171 30332: contig of 1162 bp in length
* 30333 30412: gap of unknown length
* 30413 31156: contig of 744 bp in length
* 31157 31236: gap of unknown length
* 31237 31850: contig of 614 bp in length
* 31851 31930: gap of unknown length
* 31931 33222: contig of 1352 bp in length
* 33223 33422: gap of unknown length
* 33423 34422: contig of 1120 bp in length
* 34423 35539: gap of unknown length
* 35539 36519: contig of 2037 bp in length
* 36520 37591: gap of unknown length
* 37592 37671: contig of 972 bp in length
* 37672 39179: gap of unknown length
* 39180 39260: contig of 1508 bp in length
* 39260 40565: gap of unknown length
* 40566 40645: contig of 1306 bp in length
* 40646 41733: gap of unknown length
* 41733 41812: contig of 1087 bp in length
* 41813 43094: gap of unknown length
* 43095 43174: contig of 1282 bp in length
* 43175 45242: gap of unknown length
* 45243 45322: contig of 2068 bp in length
* 45323 46265: gap of unknown length
* 46266 46345: contig of 943 bp in length
* 46346 48317: gap of unknown length
* 48318 48397: contig of 1972 bp in length
* 48398 49996: gap of unknown length
* 49997 50076: contig of 1599 bp in length
* 50077 51343: gap of unknown length
* 51344 51423: contig of 1267 bp in length
* 51424 53538: gap of unknown length
* 53539 53618: contig of 2115 bp in length
* 53619 56731: gap of unknown length
* 56732 56811: contig of 3113 bp in length
* 56812 58914: gap of unknown length
* 58915 58994: contig of 2103 bp in length
* 58995 60245: gap of unknown length
* 60246 60325: contig of 1251 bp in length
* 60326 62640: gap of unknown length
* 62641 62720: contig of 2315 bp in length
* 62721 65653: gap of unknown length
* 65654 65733: contig of 2933 bp in length
* 65734 67353: gap of unknown length
* 67354 74733: contig of 1620 bp in length
* 74734 70394: gap of unknown length
* 70395 70474: contig of 2961 bp in length
* 70475 73761: gap of unknown length
* 73762 73841: contig of 3287 bp in length
* 73842 76601: gap of unknown length
* 76602 80374: contig of 2760 bp in length
* 80375 80454: gap of unknown length
* 80455 86559: contig of 3633 bp in length
* 86559 86559: contig of 6105 bp in length

```

* 86560 86639: gap of unknown length
* 86640 90202: contig of 3563 bp in length
* 90203 90282: gap of unknown length
* 90283 94550: contig of 4268 bp in length
* 94551 94630: gap of unknown length
* 94631 98832: contig of 4202 bp in length
* 98833 98912: gap of unknown length
* 102679 102679: contig of 3767 bp in length
* 102680 102759: gap of unknown length
* 102760 102760: contig of 4427 bp in length
* 107187 107266: gap of unknown length
* 107267 114485: contig of 7219 bp in length
* 114486 114565: gap of unknown length
* 114566 115437: contig of 872 bp in length
* 115438 115517: gap of unknown length
* 115518 116166: contig of 649 bp in length
* 116167 116246: gap of unknown length
* 116247 117050: contig of 804 bp in length
* 117051 117130: gap of unknown length
* 117131 117727: contig of 597 bp in length
* 117728 117807: gap of unknown length
* 117808 118625: contig of 818 bp in length
* 118626 118705: gap of unknown length
* 118706 119280: contig of 575 bp in length
* 119281 119360: gap of unknown length
* 119361 120042: contig of 682 bp in length
* 120043 120122: gap of unknown length
* 120123 120575: contig of 453 bp in length
* 120576 120655: gap of unknown length
* 120656 121326: contig of 671 bp in length
* 121327 121406: gap of unknown length
* 121407 122063: contig of 657 bp in length
* 122064 122143: gap of unknown length
* 122144 122652: contig of 509 bp in length
* 122653 122732: gap of unknown length
* 122733 123492: contig of 760 bp in length

```

Query Match 11.1%; Score 35.6; DB 2; Length 132635;
 Best Local Similarity 51.2%; Pred. No. 4.6;
 Matches 83; Conservative 0; Mismatches 79; Indels 0; Caps 0;

```

QY 154 CAATCTACCTCCACCAACCCAGCTTGTATCTGCTTACTGATGATCACCAGATTGTGC 213
   || || || || || || || || || || || || || || || || || || || || ||
Db 21246 CATTCTCTACTTAAAGGACGACCAAGTAATATCCGTTCTCATTCGTAAGATGTAA 21187
   || || || || || || || || || || || || || || || || || || || || ||
QY 214 TCATACGATGTCGCGATTTGCTCTTCTCTCTAGATGTTCTCCGCGATGCTTTATTA 273
   || || || || || || || || || || || || || || || || || || || || ||
Db 21186 TCACGAGAAATACCAACAAATGAGATTCTCTATCAACATTTTGGGGTGGCAGCTAATTA 21127
   || || || || || || || || || || || || || || || || || || || || ||
QY 274 GAGAAAGTTGGTCAGCATGATCTCTGCGAGTGTCTAGCTGA 315
   || || || || || || || || || || || || || || || || || || || || ||
Db 21126 GCGATCGTTGATCTATCCGACGCTACGCTGCTTGAGAGA 21085
   || || || || || || || || || || || || || || || || || || || || ||

```

RESULT 15
 AC020006 156905 bp DNA linear HTG 03-JAN-2000
 LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
 AC020006
 VERSION AC020006.1 GI:6664891
 KEYWORDS HTG: HTGS, PHASE2.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 156905)
 AUTHORS Adams, M. and Venier, J.C.
 TITLE Direct Submission
 JOURNAL Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA

COMMENT
 For more information on this record e-mail to fly@celera.com.

* NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 location/Qualifiers
 1..156905
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"

BASE COUNT 44447 a 34949 c 34561 g 42948 t
 ORIGIN

Query Match 11.1%; Score 35.6; DB 2; Length 156905;
 Best Local Similarity 51.2%; Pred. No. 4.7;
 Matches 83; Conservative 0; Mismatches 79; Indels 0; Caps 0;

```

QY 154 CAATCTACCTCCACCAACCCAGCTTGTATCTGCTTACTGATGATCACCAGATTGTGC 213
   || || || || || || || || || || || || || || || || || || || || ||
Db 126540 CATTCTCTACTTAAAGGACGACCAAGTAATATCCGTTCTCATTCGTAAGATGTAA 126599
   || || || || || || || || || || || || || || || || || || || || ||
QY 214 TCATACGATGTCGCGATTTGCTCTTCTCTCTAGATGTTCTCCGCGATGCTTTATTA 273
   || || || || || || || || || || || || || || || || || || || || ||
Db 126600 TCACGAGAAATACCAACAAATGAGATTCTCTATCAACATTTTGGGGTGGCAGCTAATTA 126659
   || || || || || || || || || || || || || || || || || || || || ||
QY 274 GAGAAAGTTGGTCAGCATGATCTCTGCGAGTGTCTAGCTGA 315
   || || || || || || || || || || || || || || || || || || || || ||
Db 126660 GCGATCGTTGATCTATCCGACGCTACGCTGCTTGAGAGA 126701
   || || || || || || || || || || || || || || || || || || || || ||

```

Search completed: August 18, 2003, 13:57:18
 Job time: 2031 secs

